

Métodos basados en Independencia Condicional

Herramientas Econométricas
Prof. Ricardo Pasquini

Noviembre 2020

Sesgo en diferencias post-tratamiento

Anteriormente, obtuvimos una expresión para el resultado de tomar la diferencia ex-post entre afectados por una intervención o programa de los no afectados. Esta diferencia era igual al Tratamiento sobre los Tratados más un sesgo de selección:

$$E[Y_i | T_i = 1] - E[Y_i | T_i = 0] = \text{ATT} + \text{Preexistent bias}$$

Sesgo en diferencias post-tratamiento

$$E[Y_i \mid T_i = 1] - E[Y_i \mid T_i = 0] =$$

$$E[Y_{1i} \mid T_i = 1] - E[Y_{0i} \mid T_i = 1] + E[Y_{0i} \mid T_i = 1] - E[Y_{0i} \mid T_i = 0]$$

CIA - Selection on Observables

Ahora la innovación: la CIA afirma que, condicionalmente a una serie de características observadas X_i , los resultados potenciales son independientes del tratamiento (por ejemplo, la misma probabilidad de lograr un resultado de alto o bajo potencial si se tomó el tratamiento o no))

$$Y_{0i}, Y_{1i} \prod D_i \mid X_i$$

O de forma equivalente

$$E[Y_{0i} \mid X_i, T_i = 1] = E[Y_{0i} \mid X_i, T_i = 0]$$

CIA - Selection on Observables

Por lo tanto, el sesgo de selección desaparece y la diferencia entre participantes y no participantes dará como resultado el efecto del tratamiento:

$$E[Y_i | X_i, T_i = 1] - E[Y_i | X_i, T_i = 0] = E[Y_{1i} - Y_{0i} | T_i = 1]$$

Matching

Como podemos recuperar el ATT si se cumple CIA? :

$$\delta_{ATT} = E[Y_{1i} - Y_{0i} \mid T_i = 1]$$

$$= E[E[Y_i \mid X_i, T_i = 1] - E[Y_i \mid X_i, T_i = 0] \mid T_i = 1]$$

X CIA

$$E[Y_{0i} \mid T_i = 0] = E[Y_{0i} \mid T_i = 1]$$

Matching

Por lo tanto

$$\delta_{ATT} = E[E[Y_i \mid X_i, T_i = 1] - E[Y_i \mid X_i, T_i = 0] \mid T_i = 1]$$

$$= E[\delta_X \mid T_i = 1]$$

donde

$$\delta_X = E_Y[y_{1i} \mid X_i, T_i = 1] - E_Y[y_{0i} \mid X_i, T_i = 0]$$

Matching. El estimador

$$ATT = E[Y_{1i} \mid T_i = 1] - E[Y_{0i} \mid T_i = 1] = \sum_x \delta_x P(X_i = x \mid T_i = 1)$$

CASO: Dale y Krueger 2002

ESTIMATING THE PAYOFF TO ATTENDING A MORE SELECTIVE COLLEGE: AN APPLICATION OF SELECTION ON OBSERVABLES AND UNOBSERVABLES*

STACY BERG DALE AND ALAN B. KRUEGER

Estimates of the effect of college selectivity on earnings may be biased because elite colleges admit students, in part, based on characteristics that are related to future earnings. We matched students who applied to, and were accepted by, similar colleges to try to eliminate this bias. Using the College and Beyond data set and National Longitudinal Survey of the High School Class of 1972, we find that students who attended more selective colleges earned about the same as students of seemingly comparable ability who attended less selective schools. Children from low-income families, however, earned more if they attended selective colleges.

Figure 1: 1565209052179

CASO: Dale y Krueger 2002

"Past studies have found that students who attended colleges with higher average SAT scores or higher tuition tend to have higher earnings when they are observed in the labor market. Attending a college with a 100 point higher average SAT is associated with 3 to 7 percent higher earnings later in life (see, e.g., Kane [1998]). As Kane notes, an obvious concern with this conclusion is that students who attend more elite colleges may have greater earnings capacity regardless of where they attend school."

CASO: Dale y Krueger 2002 (simplificado)

TABLE 2.1
The college matching matrix

Applicant group	Student	Private			Public			1996 earnings
		Ivy	Leafy	Smart	All State	Tall State	Altered State	
A	1		Reject	Admit		Admit		110,000
	2		Reject	Admit		Admit		100,000
	3		Reject	Admit		Admit		110,000
B	4	Admit			Admit		Admit	60,000
	5	Admit			Admit		Admit	30,000
C	6		Admit					115,000
	7		Admit					75,000
D	8	Reject			Admit	Admit		90,000
	9	Reject			Admit	Admit		60,000

Note: Enrollment decisions are highlighted in gray.

CASO: Dale y Krueger 2002 (simplificado)

Ver que:

- ▶ La diferencia de privada vs publica me da aprox 20,000 (92 vs 72,5 en promedio)
- ▶ La diferencia si matcheamos por grupo y ponderamos es:

$$\frac{3}{5}(-5000) + \frac{2}{5}30,000 = 9000$$

Al matchear tengo que descartar los grupos C y D

CASO: Dale y Krueger 2002 (simplificado)

Traducir el matching en un estimador de regresión

CASO: Dale y Krueger 2002 (simplificado)

TABLE I
ILLUSTRATION OF HOW MATCHED-APPLICANT GROUPS WERE CONSTRUCTED

Student	Matched-applicant group	Student applications to college							
		Application 1		Application 2		Application 3		Application 4	
		School average	SAT	School admissions decision	School average	SAT	School admissions decision	School average	SAT
Student A	1	1280	Reject	1226	Accept [*]	1215	Accept	na	na
Student B	1	1280	Reject	1226	Accept	1215	Accept [*]	na	na
Student C	2	1360	Accept	1310	Reject	1270	Accept [*]	1155	Accept
Student D	2	1355	Accept	1316	Reject	1270	Accept [*]	1160	Accept
Student E	2	1370	Accept [*]	1316	Reject	1260	Accept	1150	Accept
Student F	Excluded	1180	Accept [*]	na	na	na	na	na	na
Student G	Excluded	1180	Accept [*]	na	na	na	na	na	na
Student H	3	1360	Accept	1308	Accept [*]	1260	Accept	1160	Accept
Student I	3	1370	Accept [*]	1311	Accept	1255	Accept	1155	Accept
Student J	3	1350	Accept	1316	Accept [*]	1265	Accept	1155	Accept
Student K	4	1245	Reject	1217	Reject	1180	Accept [*]	na	na
Student L	4	1235	Reject	1209	Reject	1180	Accept [*]	na	na
Student M	5	1140	Accept	1055	Accept [*]	na	na	na	na
Student N	5	1145	Accept [*]	1060	Accept	na	na	na	na
Student O	No match	1370	Reject	1038	Accept [*]	na	na	na	na

* Denotes school attended.

na = did not report submitting application.

The data shown on this table represent hypothetical students. Students F and G would be excluded from the matched-applicant subsample because they applied to only one school (the school they attended). Student O would be excluded because no other student applied to an equivalent set of institutions.

CASO: Dale y Krueger 2002 (simplificado)

- Un modelo de regresión va a devolver una estimación similar a la de matching:

$$\log(earnings) = \alpha_i + \beta P_i + \beta SAT_i + \varepsilon_i$$

- La siguiente tabla (Columnas 1-3) muestran esta especificación básica

CASO: Dale y Krueger 2002 (simplificado)

TABLE 2.2
Private school effects: Barron's matches

[From Mastering Metrics: The Path from Cause to Effect. © 2013 Princeton University Press. Used by permission. All rights reserved.]

	No selection controls			Selection controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	.135 (.055)	.095 (.052)	.086 (.034)	.007 (.038)	.003 (.039)	.013 (.025)
Own SAT score ÷ 100		.048 (.009)	.016 (.007)		.033 (.007)	.001 (.007)
Log parental income			.219 (.022)			.190 (.023)
Female				−.403 (.018)		−.395 (.021)
Black				.005 (.041)		−.040 (.042)
Hispanic				.062 (.072)		.032 (.070)
Asian				.170 (.074)		.145 (.068)
Other/missing race				−.074 (.157)		−.079 (.156)
High school top 10%				.095 (.027)		.082 (.028)
High school rank missing				.019 (.033)		.015 (.037)
Athlete				.123 (.025)		.115 (.027)
Selectivity-group dummies	No	No	No	Yes	Yes	Yes

Notes: This table reports estimates of the effect of attending a private college or university on earnings. Each column reports coefficients from a regression of log earnings on a dummy for attending a private institution and controls. The results in columns (4)–(6) are

Control for Covariates Using the Propensity Score

Propensity score matching (PSM) constructs a statistical comparison group that is based on a model of the probability of participating in the treatment, using observed characteristics. The idea is to match each individual in the treatment group with a individual in the control group with a similar probability of taking up the program.

Propensity Score Matching (PSM)

This involves two assumptions: CIA: No program take up due to unobserved factors. Self-selection bias is eliminated. In this, the PSM technique is therefore similar to simple Regression and Matching. Presence of a common support or overlap condition, which guarantees that I will be able to find a “similar” individual in the control group.

$$0 < P(T_i = 1|X_i) < 1$$

if

$$P(T_i = 1|X_i) = 1$$

, all of them would be takers...

Propensity Score Matching (PSM) - Common Support Example

Figure 4.1 Example of Common Support

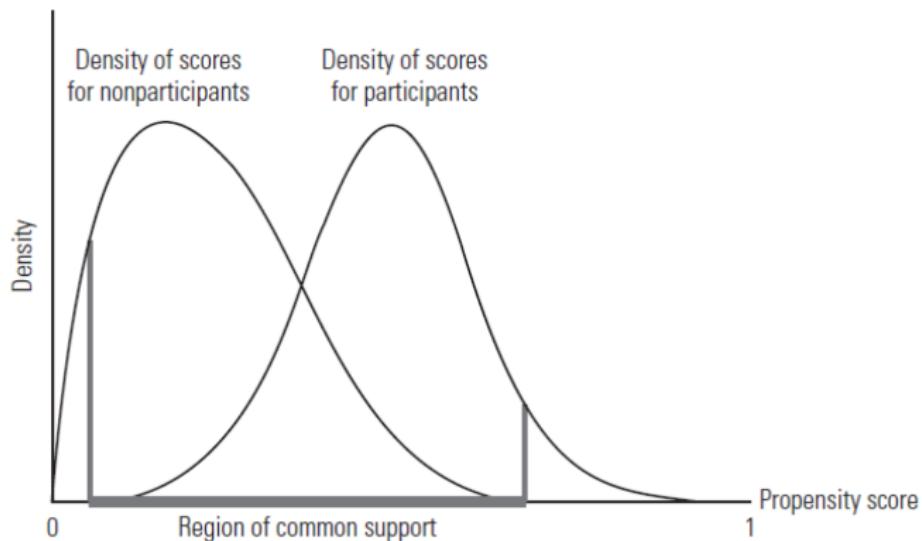


Figure 6: 1565205953462

Propensity Score Matching (PSM)

In practice, estimation in three steps: - First, $p(T_i = 1|X_i)$ is estimated using some kind of parametric model, say, Logit or Probit.

- ▶ Second, different approaches can be used to match participants and nonparticipants on the basis of $p(T_i = 1|X_i)$
- ▶ Third, estimates of the TOT are computed by evaluating the differences.

$$\text{TOT}_{PSM} = \frac{1}{N_T} \left[\sum_{i \in T} Y_i^T - \sum_{i \in T} w(i, j) Y_i^C \right]$$

where $w(i, j)$ are weights when there are more than one control for each treated observation. An alternative is to use a regression and the $p(T_i = 1|X_i)$ as weights.

Propensity Score Matching (PSM)

Matching participants and non-participants:

- ▶ Nearest-neighbor: just take the n nearest neighbors on the basis of $p(T_i = 1|X_i)$.
- ▶ Caliper or radius matching: there might not be long distance neighbors, then impose a radius or threshold.
- ▶ Stratification or interval matching. Partitionates the common support into strata (or intervals) and calculates the program's impact within each interval.
- ▶ Kernel Matching: Takes into account all possible matches and weights them according to distance.

Propensity Score Matching (PSM)

- ▶ Bootstrap Calculation of Standard Errors
- ▶ Testing for the presence of selection bias due to unobserved factors: The Sargan-Wu-Hausman Test.
- ▶ Using propensity scores as weights in a regression approach.
Hirano, Imbens, and Ridder (2003) Estimate:
$$Y_i = \alpha + \beta T_i + \rho X_i + \varepsilon_i$$
 with weights of 1 for participants and
weights of $\hat{P}(X)/(1 - \hat{P}(X))$ for the control observations.

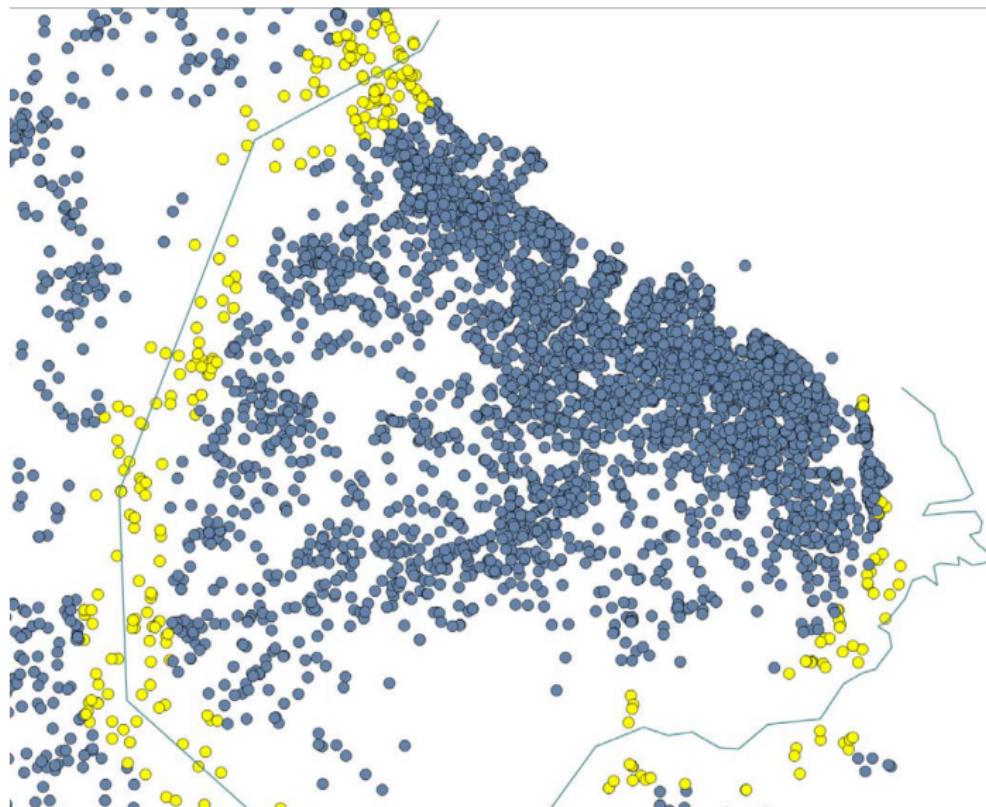
Caso: Efectos ley de alquileres

El 8 de Septiembre de 2017 la legislatura de la Ciudad de Buenos Aires dictaminó la ley 5859, que entró en vigencia el 16 de septiembre (más info en las noticias: [Aquí](#), [aquí](#)). Esta ley fue sancionada con el objetivo de beneficiar a los inquilinos. La ley estableció, en primer lugar, que la comisión inmobiliaria no sería pagada por los inquilinos, sino por los dueños. Esto modificó lo que era la práctica habitual, en la cual se cargaba al inquilino con el costo. Esto implicaría una reducción en el costo de entrada a un contrato de alquiler para los inquilinos, ya que estos en muchos casos solían pagar un mes, o más en términos de comisión. Además, la ley estableció un máximo para esta comisión, correspondiente a un mes extra (el equivalente a 4,15% -1 mes sobre 24- del valor total del contrato)

Caso: Efectos ley de alquileres

Estrategia de Identificación: Enfoque cuasi-experimental.

Matching con datos post-intervención>



Caso: Efectos ley de alquileres



Ejemplo Matching results 1

Proprietà - Capital Federal - Nuñez - Local - Alquiler

ALQUILER

Local comercial en Venta / Alquiler - Av. Cabildo 4700 - Nuñez - Local Comercial En Venta / Alquiler - Av. Cabildo 4700 - Nuñez - \$ 115.000
Locales en Alquiler en Nuñez, Capital Federal, Argentina



Proprietà - B.A. G.B.A. Zona Norte - Nueve Locas - Florida - Local - Alquiler

ALQUILER

Único Local Comercial, Apto Todo Destino sobre Av. Maipú - Florida - Maipú Al 800 - \$ 150.000
Locales en Alquiler en Florida, Vicente López, Bs.As. G.B.A. Zona Norte, Argentina

Detalle Mapa StreetView Fotos en galería

► Descripción

PLANTA BAJA/200m2 cubiertos en planta baja totalmente libres. Pisos acceso vehicular con puerta especial concreta (en 3 planos) para facilitar la carga y descarga de mercaderías, así como, el acceso de automóviles, motos y camionetas al interior del salón ya que cuenta con una altura de techos superior a 3mts. Esta

Contactar al anfitrión: Carlos Castillo Propiedades
Número:

Caso: Efectos ley de alquileres

Ejemplo Matching results 2

Properati · Bs.As. G.B.A. Zona Norte · Vicente López · Florida · Departamento · Alquiler ·

Depto 2 Amb Exc Vista a 2/C de Maipu - Florida - Melo Al 1600 - \$ 7.000
Departamentos en Alquiler en Florida, Vicente López, Bs.As. G.B.A. Zona Norte, Argentina

Este aviso ha caducado

Properati · Capital Federal · Nuñez · Departamento · Alquiler ·

Unico dos ambientes en Torre Astor con cochera y baulera. - Nuñez - Pico Al 2300 Astor Núñez - \$ 16.500
Departamentos en Alquiler en Nuñez, Capital Federal, Argentina

Este aviso ha caducado

Figure 9: 1565207368637



Ejemplo Matching results 3